An Engineering Approach to Nonlinear Estimation

by J. H. Park, Jr.

Abstract

Parameter identification in nonlinear systems is very similar to the estimation of certain signal parameters. In both cases the problem is to estimate a time invariant parameter that cannot be extracted from the data by a linear operation. The data is observed in the presence of additive noise for a finite time duration T. For this nonlinear estimation problem a solution is desired that is readily realizable by physical systems.

In the proposed solution to this problem the signal processing equipment is restricted to be linear time invariant circuits "read" sometime in the observation interval, followed by a quantizer and digital computer. The various impulse responses of these linear circuits are designed to be orthogonal over the interval $[0, T]$; hence, they span a subspace, S_s, of function space. The computer operates on the components of the data in this subspace. The computer operations are restricted to be linear or multiplication/division or square/square root. It is shown how maximum likelihood estimators and Bayes estimators with a quadratic cost function for a large class of parameters can be determined exactly or closely approximated using the above techniques. Several examples are worked out in detail such as the estimation of the width of a pulse and first order [illegible]

*Supported by NASA Grant NsG 712

U-Minnesota

Code 1
CR-75892
pages-22

HC 1.00
MF .50

N66 29569

Cat. 19

I. Introduction

The observed time functions or signals in communication and control problems can be represented in a variety of ways. In communication work one usually considers the observation of a single function. This function is often treated as a vector in a linear vector space. For example, if it is band-limited the sample values are the vector components and the cardinal function is the tacitly assumed interpolating or coordinate function. With random signals it is convenient to use the Karhunen-Loeve expansion to obtain vector components (i.e., expansion coefficients) that are uncorrelated. In control problems one treats "state variables" and the observables are a set of state variables or a state vector which is a (finite) vector time function. There is no reason why each of the components of this state vector cannot be treated as the single observation of the communication problem resulting in a representation of each state vector as a vector with time invariant components.

In a practical engineering situation one may wish to truncate such a vector and operate on the resulting projection. For example, it is extremely inconvenient to use all the sample values of a band-limited signal. One usually restricts the observation to a finite time interval hence a finite number of samples. Any parameter that may be of interest is assumed to be embedded in one or more of these time invariant coefficients.

In general a nonlinear operation on a set of these coefficients is required to obtain an estimate of the desired parameter. The form of an estimator has now been specified, it consists of linear operations to obtain the time invariant coefficients and nonlinear operations on the resultant set of numbers to obtain the parameter estimate. That is, a linear analog operation followed by nonlinear digital operations, a potentially easy physical realization. One must consider how close such estimates are to the optimum.

In this paper it is assumed that a single function of time (one state variable) is observed with known structure and an unknown parameter and unknown amplitude in the presence of additive Gaussian noise. That is, $v(t) = As(t,\alpha) + n(t)$ is observed. An observation time T is assumed and a set of linear operators $L_1, L_2, \ldots L_n$ with corresponding time functions $\ell_1(t), \ell_2(t), \ldots \ell_n(t)$ orthogonal on the T interval are used. The noise is assumed to be zero mean with a flat (white) spectrum. The resultant set of coefficients are therefore Gaussian, time independent, and statistically independent. An appropriate nonlinear operation is then used which gives a near optimum estimate of α independent of A. [illegible] of this type [illegible] that the variance of the [illegible] based on the [illegible] of $v(t)$ [illegible] g_j [illegible] the variance based on all of $v(t)$.

II. The Near-Optimum Estimators

Optimum estimators are conventionally defined in terms of risk theory utilizing a preassigned cost function. If a quadratic cost function is used the optimum (Bayes) estimate is just the expected value of the parameter with respect to the posterior statistics. In the case of maximum likelihood estimation (Bayes estimate with a simple cost function) the estimate is that value of the parameter corresponding to the peak of the posterior statistics. In either case knowledge of the posterior statistics is sufficient to define the optimum estimator.

In the present context $a_1, a_2, \ldots a_n$ are defined as the vector components of $s(t,\alpha)$ in S_s, that is, $L_k\left[s(t,\alpha)\right] = a_k$. Similarly $b_1, b_2, \ldots, b_n$ are those components associated with $n(t)$ and $c_1, c_2, \ldots, c_n$ with $v(t)$. $\bar{a}$, $\bar{b}$, and $\bar{c}$ represents the corresponding ordered n-tuples of these coefficients. Hence $\bar{c} = A\bar{a} + \bar{b}$. From the assumptions on $n(t)$ the b_k's are zero mean, Gaussian and statistically independent. The posterior density distribution of α, $p(\alpha/\bar{c})$ is therefore

$$p(\alpha/\bar{c}) = \left[1/K\right] \int_A p(\bar{c}/\alpha, A)\; p(\alpha, A)\, dA \qquad \text{II-1}$$

where

$$K = \int_A \int_\alpha p(\bar{c}/\alpha, A)\; p(\alpha, A)\, d\alpha\, dA \qquad \text{II-2}$$

and is independent of α. $p(\alpha, A)$ is the a priori probability density of α and A. $p(\bar{c}/\alpha, A)$ is obtained from the known noise statistics. That is

$$p(\bar{c}/\alpha, A) = \left[\text{constant}\right] \exp\left\{\sum_{i=1}^{n} \left[c_i - Aa_i(\alpha)\right]^2 / 2\sigma_i^2\right\} \qquad \text{II-3}$$

where the α dependence of the a_k's has been indicated. If the L_i's are normalized then the σ_i's are identical. The difficulty in obtaining optimum estimates of α from II-3 and II-1 is due to the nonlinear dependence of the a_i's on α.

Engineering judgement has already been necessary in the choice of the L's. Further engineering judgement is now required to find a set of nonlinear transformations (coordinate change if you like) on the a's $f_1(\bar{a})$, $f_2(\bar{a})$, ... $f_n(\bar{a})$ such that $f_i(\bar{a}) = \alpha$ and the set of functions $\{f_i\}$ has a corresponding inverse set $\{g_i\}$ such that $g(f(\bar{a})) = \bar{a}$. The first condition is almost tantamount to guessing the optimum estimator. The second assumption will insure us that no information is lost in the transformation, that $p(\alpha/\bar{c}) = p(\alpha/f(\bar{c}))$. With these assumptions and II-1 the posterior statistics of α become

$$p\left[\alpha/f(\bar{c})\right] = (1/K) \int_A p(f(\bar{c})/\alpha, A)\ p(\alpha, A)\, dA \qquad \text{II-4}$$

As will be seen in the examples $p(\alpha/f(\bar{c}))$ often has a [illegible] expected value that depends only on $f_i(\bar{c})$. In this case the near-optimum estimator consists of the n linear operators L_1, L_2, ... L_n and the nonlinear operator f_i.

The variable $\tilde{c}$ has Gaussian statistics (non-central) hence $p(f_1(\tilde{c})/\alpha)$ can be obtained for various nonlinear functions f_1. Hence the statistics of the estimate may be readily evaluated. Determination of whether the estimate is biased or not ($E(f_1(\tilde{c})) = \alpha$?) is easily done as well as the evaluation of the variance of the estimate. At this point the utility of this approach might be questioned based on the required assumptions and dependence on engineering judgement. The applications of the next section should quench any such apprehensions.

III. Applications

A. Pulse Width

The term "pulse width" is unfortunately an ambiguous term. An unambiguous definition that satisfies our intuition might be "radius of gyration" (or square root of the second moment about the mean). In this section the square of the pulse width will be considered as the parameter of interest and defined in such a way as to fit in the context of section II. It is then shown that for three common pulse shapes that this definition gives a width almost equal to the radius of gyration. The three pulse shapes considered are tabulated below.

The [illegible] also be assumed to be [illegible] width [illegible] [illegible]

Table 1

Pulse Shape under Consideration

Shape	Functional Form	Radius of Gyration
Square	$A\left[u(t) - u(t-b)\right]$	$b/\sqrt{3}$
Triangular	$A\left\{(t/b)u(t) + \left[2(b-t)/b\right]u(t-b) - \left[(2b-t)/b\right]u(t-2b)\right\}$	$b/\sqrt{6}$
Gaussian	$A\exp(-t^2/2b^2)$	b

for the Gaussian pulse. The two linear operators that will be used have the following orthonormal corresponding time functions

$$\ell_1(t) = \left[2/(3/\pi)^{1/2}\right] t^2 \exp(-t^2/2)$$
$$\ell_2(t) = \left[2/3/\pi\right]^{1/2} \left[3/2 - t^2\right] \exp(-t^2/2) \qquad \text{III-1}$$

The nonlinear functions f_1 and f_2 are

$$f_1(a) = a_1/a_2$$
$$f_2(a) = (a_1^2 + a_2^2)^{1/2} \qquad \text{III-2}$$

with corresponding inverses

$$a_1 = f_1 f_2/(1 + f_1^2)^{1/2}$$
$$a_2 = f_2/(1 + f_1^2)^{1/2} . \qquad \text{III-3}$$

The square of the pulse width is defined now as $f_1(a) = W^2$. Table 2 gives the coefficients a_1 and a_2 [illegible] the three shapes and also compares W with [illegible] The results given are [illegible] they are the first terms in [illegible]

Table 2

Pulse shape	a_1	a_2	$W=(\frac{a_1}{a_2})^{1/2}$	Radius of gyration
Square	$Ab^3(2/9)(12/\sqrt{\pi})^{1/2}$	$Ab(b/\sqrt{\pi})^{1/2}$	(.56)b	(.58)b
Triangle	$Ab^3(1/3)(1/3/\pi)^{1/2}$	$Ab(3/2)(2/3/\pi)^{1/2}$	(.40)b	(.41)b
Gaussian	$2b^2A(2/\pi/3)^{1/2}$	$3Ab(\sqrt{\pi}/3)^{1/2}$	(.97)b	(1.0)b

The definition of pulse width as $(a_1/a_2)^{1/2}$ appears to satisfy our intuition as long as the width is small compared to the "variance" of the ℓ_1 and ℓ_2 functions (assumed one in this example, see III-1).

The joint probability density of c_1 and c_2 given a_1 and a_2 is

$$p(c_1,c_2/a_1,a_2) = (1/2\pi\sigma^2)\exp\left\{-\frac{1}{2\sigma^2}\left[(c_1-a_1)^2+(c_2-a_2)^2\right]\right\} \quad \text{III-4}$$

The transformation from the variables c_1,c_2 to $f_1(\bar{c})$, $f_2(\bar{c})$ is easily performed resulting in

$$p(f_1,f_2/a_1,a_2) = (1/2\pi\sigma^2)\left[f_2/(1+f_1^2)\right]\exp\left[-\frac{1}{2\sigma^2}(f_2^2+a_1^2+a_2^2)\right]$$
$$\times\left\{\exp\left[\frac{f_2(a_1f_1+a_2)}{\sigma^2\sqrt{(1+f_1^2)}}\right]+\exp\left[-\frac{f_2(a_1f_1+a_2)}{\sigma^2\sqrt{(1+f_1^2)}}\right]\right\} \quad \text{III-5}$$

It is now desirable to obtain this conditional density in terms of the parameters W and A rather than a_1 and a_2. From Table 2 it can be seen that A is proportional to $(a_2^3/a_1)^{1/2}$ hence, for mathematical simplicity A is so defined. That is

$$A^2 = a_2^{\,3}/a_1$$
$$W^2 = a_1/a_2 \qquad \text{III-6}$$

with the inverse relations

$$a_2 = AW$$
$$a_1 = AW^3 \,. \qquad \text{III-7}$$

All the computations to obtain posterior statistics are now independent of pulse shape.

If III-7 is used in III-5 the following conditional density is obtained after a little rearranging of terms.

$$p(f_1, f_2/A, W) = (1/\pi\sigma^2)\left(\frac{f_2}{1+f_1^{\,2}}\right) \exp\left[-(f_2^{\,2}+A^2W^2+A^2W^6)/2\sigma^2\right]$$

$$\times \cosh\left[\frac{AWf_2(W^2f_1+1)}{\sigma^2(1+f_1^{\,2})^{1/2}}\right] . \qquad \text{III-8}$$

The desired posterior statistics are obtained by multiplying III-8 by $p(A,W)$ and integrating over A. A priori we assume $A > 0$ and from III-2 $f_2 > 0$. Hence for high S/N (that is AWf_2/σ^2 large compared to 1) cosh is closely approximated by 1/2 exp in III-8. The expression for the posterior statistics for high signal to noise is therefore,

$$p(W/f_1 \cdot f_2) = (1/2\pi\sigma^2)\left(\frac{f_2}{1+f_1^{\,2}}\right) \exp(-f_2^{\,2}/2\sigma^2) \qquad \text{III-9}$$

$$\times \int_A \exp\left\{-\left[A^2W^2(1+W^4) - \frac{2AWf_2(W^2f_1+1)}{(1+f_1^{\,2})^{1/2}}\right]/2\sigma^2\right\} p(A,W)\,dA \,.$$

Completing the square in A in the exponential under the integral sign and assuming $p(A,W) = p(A)p(W)$ yields

$$p(W/f_1,f_2) = \left[p(W)/2\pi\sigma^2\right]\left(\frac{f_2}{1+f_1^2}\right) \exp\left\{-\frac{f_2^2}{2\sigma^2}\left[1-\frac{(W^2f_1+1)^2}{(1+W^4)(1+f_1^2)}\right]\right\}$$

$$\times\int_A \exp\left\{-\left[AW(1+W^4)^{1/2}-\frac{f_2(W^2f_1+1)}{(1+W^4)^{1/2}(1+f_1^2)^{1/2}}\right]^2\Big/2\sigma^2\right\}p(A)\,dA$$

III-10

If $p(A)$ is "flat" compared to the exponential (that is a priori knowledge is vague) then the integral evaluates easily and we finally obtain

$$p(W/f_1,f_2) = \text{constant}\cdot p(W)\cdot\left\{W^2(1+W^4)\left[1+\frac{(f_1-W^2)^2}{(1+W^2f_1)^2}\right]\right\}^{-1/2}$$

$$\exp\left\{-\frac{f_2^2}{2\sigma^2}\left[1-\frac{1}{1+\frac{(f_1-W^2)^2}{(1+W^2f_1)^2}}\right]\right\}.$$

III-11

For the above W to satisfy our intuitive notion concerning pulse width we required that $W < 1$. Let us further assume that $f_1 < 1$ so that III-11 is approximated by

$$p(W/f_1,f_2) \cong \text{constant}\cdot\left(\frac{1}{W}\right)\cdot\frac{1}{\sqrt{2\pi}(\sigma/f_2)}\exp\left[-\frac{(f_1-W^2)^2}{2\sigma^2/f_2^2}\right] \qquad \text{III-12}$$

That is, for high S/N and W bounded away from zero W^2 has Gaussian posterior statistics with mean f_1 and variance σ^2/f_2^2.

The estimate of W^2 is therefore f_1 for high S/N and is unbiased. As the signal to noise decreases the estimate will be something closer to zero than f_1 and also becomes biased.

B. AM and FM Detection

Consider a narrowband signal together with narrowband noise, the signal may be amplitude modulated or frequency (or phase) modulated in both cases we assume the carrier frequency is known. We will now determine the optimum estimate of instantaneous amplitude (unknown phase) or frequency (unknown amplitude). Optimum detectors of am and fm should take into account the second order statistics as well as the first order statistics. In this treatment we assume Weiner type smoothing before and/or after the estimator and determine the estimator from first order statistics.

The signal is s(t),

$$s(t) = A(t) \cos\left[\omega_c t + \theta + \varphi(t)\right] \qquad \text{III-13}$$

$$= A(t) \cos \omega_c t \cos\left[\theta + \varphi(t)\right] - A(t) \sin \omega_c t \sin\left[\theta + \varphi(t)\right] .$$

The noise n(t) can be represented by an in-phase and quadrature component, i.e.,

$$n(t) = n_1(t) \cos \omega_c t - n_2(t) \sin \omega_c t \qquad \text{III-14}$$

where $n_1(t)$, and $n_2(t)$ are independent [illegible] with the [illegible] first order statistics as n(t). The [illegible] multiplying by $\cos \omega_c t$ and [illegible] number of periods of the carrier (low pass filtering gives the same result). Hence the vector components of interest are

$$c_1 = A\cos(\varphi + \theta) + n_1(t)$$
$$c_2 = A\sin(\varphi + \theta) + n_2(t) \qquad \text{III-15}$$

Since n_1 and n_2 are zero mean Gaussian and independent we know

$$p(c_1, c_2/A, \varphi, \theta) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{[a_1 - A\cos(\varphi+\theta)]^2 + [a_2 - A\sin(\varphi+\theta)]^2}{2\sigma^2}\right) \qquad \text{III-16}$$

The nonlinear functions f_1 and f_2 are

$$f_1 = \sqrt{(a_1^2 + a_2^2)}$$
$$f_2 = \cos^{-1}\left[a_1/(a_1^2 + a_2^2)^{1/2}\right] = \tan^{-1}(a_2/a_1) \qquad \text{III-17}$$

The statistics of f_1 and f_2 given A, φ, θ are easily found from III-16.

$$p(f_1, f_2/A, \varphi, \theta) = (A/4\pi^2\sigma^2)\exp\left[-\frac{A^2 + f_1^2 - 2Af_1\cos(f_2 - \varphi - \theta)}{2\sigma^2}\right] \qquad \text{III-18}$$

for $A > 0$ and $0 < f_2$, φ, $\theta < 2\pi$

In the case of am detection we assume $\varphi = 0$ and θ is uniformly distributed from 0 to 2π. Furthermore we assume A and θ are a priori independent hence

$$p(A/f_1, f_2) = \text{constant} \cdot p(A) \cdot (A/4\pi^2\sigma^2)\exp(-A^2/2\sigma^2)\exp(-f_1^2/2\sigma^2)\int_0^{2\pi}\exp\left(\frac{Af_1\cos(f_2 - \theta)}{\sigma^2}\right)d\theta \qquad \text{III-19}$$

This gives the well known result

$$p(A/f_1, f_2) = \text{constant} \cdot p(A) \cdot A \exp\left[-A^2/2\sigma^2\right] I_0\left(\frac{Af_1}{\sigma^2}\right) . \qquad \text{(II-20)}$$

For high signal to noise and "flat" a priori probabilities this is known to be Gaussian with mean f_1 and variance σ^2, i.e., the optimum detector is the envelope detector.

The more interesting case is when A is constant and we wish an estimate of $d\varphi/dt$ or φ. We assume φ is uniformly distributed a priori, θ is identically zero, A has a Rayleigh distribution and A and φ are independent, that is, a priori

$$p(A,\theta,\varphi) = \begin{cases} \frac{1}{2\pi}\, p(A) & 0 \le \varphi \le 2\pi,\ \theta = 0 \\ 0 & \text{otherwise} \end{cases} \qquad \text{(II-21)}$$

and

$$p(A) = \begin{cases} (A/\sigma_1^2) \exp(-A^2/2\sigma_1^2) & A > 0 \\ 0 & \text{otherwise} . \end{cases} \qquad \text{(II-22)}$$

In this case

$$p(\varphi/f_1, f_2) = \text{constant} \int_0^{\infty} A p(A) \exp\left[\frac{-A^2 + 2Af_1\cos(f_2 - \varphi)}{2\sigma^2}\right] dA \qquad \text{(II-23)}$$

[illegible]

[illegible]

$$p(\varphi/f_1, f_2) = \text{constant}\ \exp\left[\frac{\beta f_1^2 \cos^2(f_2-\varphi)}{2\sigma^4}\right] \int_0^\infty A^2 \exp\left[-\frac{\left(A - \frac{\beta f_1 \cos(f_2-\varphi)}{\sigma^2}\right)^2}{2\beta}\right] dA \qquad \text{III-24}$$

where

$$\beta = \frac{\sigma_1^2 \sigma^2}{\sigma_1^2 + \sigma^2} . \qquad \text{III-25}$$

For high signal to noise, that is $\sigma_1^2 >> \sigma^2$ and $f_1 \cos(f_2-\varphi) >> \sigma$, $\beta \cong \sigma^2$ and the A integral in evaluated by extending the lower limit to $-\infty$. Hence

$$p(\varphi/f_1, f_2) \cong \text{constant}\ \exp\left[f_1^2 \cos^2(\varphi-f_2)/2\sigma^2\right] \times \left[1 + f_1^2 \cos^2(\varphi-f_2)/\sigma^2\right] \qquad \text{III-26}$$

This is clearly symmetric about $\varphi = f_2$ and peaks at $\varphi = f_2$ hence the optimum estimate of φ is f_2 and is unbiased.

If A is known exactly prior to detection, that is $p(A) = \delta(A_0)$ then from III-23 we obtain

$$p(\varphi/f_1, f_2) = \text{constant} \times \exp(Af_1 \cos(f_2-\varphi)/\sigma^2) \qquad \text{III-27}$$

Again this is symmetric about $\varphi = f_2$ and peaks at this same value hence the optimum estimate is again f_2. The optimum [illegible] f_2 [illegible] modulation is $\dot{f}_2$ since $F(d\varphi/dt) = d\left[F(\varphi)\right]/dt$ or returning to III-17 we see that the optimum estimate of frequency modulation is

$$\frac{df_2}{dt} = \frac{[illegible]}{\sigma_1^2 + \sigma_2^2} \qquad \text{III-28}$$

This is what is known as the 'ideal detector'. The results here are based on first order statistics. It appears that, at least for high signal to noise ratios this first order optimum estimate followed by a Weiner filter (minimum mean square error filter) would be the optimum when considering second order statistics.

C. Velocity Error of a Narrow Band Process

Let x(t) be a narrowband stochastic process centered at ω_c radians/sec. x(t) may be written as

$$x(t) = A(t)\ \cos\left[\omega_c t + \varphi(t)\right] \qquad \text{III-29}$$

where A(t) and φ(t) vary slowly with time. This x(t) is observed over a time interval spanning several cycles of $\cos\omega_c t$ but over which A and φ are essentially constant and there is additive zero mean noise also narrowband. There is a desired velocity $v_d(t)$ which also comes from a narrowband position $x_d(t)$. An estimate of the velocity error $\varepsilon(t) = v_d(t) - dx(t)/dt$ is desired. We assume the actual position and velocity are close to the desired position and velocity.

Let

$$x_d(t) = B(t)\ \cos\left[\omega_c t + \theta(t)\right] \qquad \text{III-30}$$

hence

$$v_d(t) = B'(t)\ \cos\left[\omega_c t + \theta(t)\right] - B(t)\left[\omega_c + \theta'(t)\right]\sin\left[\omega_c t + \theta(t)\right]$$

$$\text{III-31}$$

Note B, θ and ω_c are known to the observer but A and φ are not. The actual velocity $v(t)$ is

$$v(t) = dx(t)/dt = A'(t) \cos\left[\omega_c t + \varphi(t)\right] - A(t)\left[\omega_c + \varphi'(t)\right]\sin\left[\omega_c t + \varphi(t)\right] \qquad \text{III-32}$$

Since A, B, φ, and θ are slowly varying $v(t)$ and $v_d(t)$ are, to a first approximation, given by

$$\begin{aligned} v(t) &\cong -A(t)\omega_c \sin\left[\omega_c t + \varphi(t)\right] \\ v_d(t) &\cong -B(t)\,\omega_c \sin\left[\omega_c t + \theta(t)\right]. \end{aligned} \qquad \text{III-33}$$

Letting $A = B + \Delta B$ and $\varphi = \theta + \Delta\theta$ and assuming ΔB and $\Delta\theta$ are small the velocity error becomes

$$\begin{aligned} \varepsilon(t) &= v_d(t) - v(t) \\ &= \omega_c \Delta\theta x_d(t) - \frac{\Delta B}{B} v_d(t) \end{aligned} \qquad \text{III-34}$$

where $\Delta\theta$ and ΔB are the unknowns.

Note we observe

$$U(t) = (B+\Delta B) \cos(\omega_c t+\theta+\Delta\theta) + n_1(t) \cos(\omega_c t+\theta) - n_2(t) \sin(\omega_c t+\theta) \qquad \text{III-35}$$

The disturbance has been written as an in-phase and quadrature component. As in section 2 we use the linear operators of multiplying by $\cos(\omega_c t+\theta)$ and $\sin(\omega_c t+\theta)$ and integrating over the observation interval. Hence we obtain the [illegible]

$$\begin{aligned} c_1 &= (B + \Delta B) \cos\Delta\theta + n_1(t) \\ c_2 &= (B + \Delta B) \sin\Delta\theta + n_2(t) \end{aligned} \qquad \text{III-36}$$

Again using f_1 and f_2 as defined in III-16 we obtain

$$p(f_1, f_2/B, \Delta B, \Delta\theta) = \left[(B + \Delta B)/4\pi^2\sigma^2\right]$$

$$\times \exp\left[-\frac{(B+\Delta B)^2 - 2f_1(B+\Delta B)\cos(f_2 - \Delta\theta) + f_1^2}{2\sigma^2}\right] \qquad \text{III-37}$$

The error is a linear combination of $\Delta\theta$ and ΔB hence we must obtain

$p(\Delta B, \Delta\theta/f_1, f_2)$.

$$p(\Delta\theta, \Delta B/f_1, f_2) = p(\Delta\theta, \Delta B)\, p(f_1, f_2/\Delta\theta, \Delta B, B) \quad . \qquad \text{III-38}$$

We now use a high signal to noise approximation together with the small error approximation, namely, we let

$$\cos(f_2 - \Delta\theta) \cong 1 - \frac{(f_2 - \Delta\theta)^2}{2}$$

and obtain from III-38 and III-37

$$p(\Delta\theta, \Delta B/f_1, f_2) \cong p(\Delta\theta, \Delta B)\, \exp\left[-\frac{(\Delta B - (B - f_1))^2}{2\sigma^2}\right] \left(\frac{1}{\sqrt{2\pi}\sigma}\right)$$

$$\times \exp\left[\frac{(f_2 - \Delta\theta)^2}{2\sigma^2/Bf_1}\right] \left(1/\sqrt{2\pi}\frac{\sigma}{\sqrt{Bf_1}}\right)$$

If a priori $p(\Delta\theta, \Delta B)$ is "flat" we see that $\Delta\theta$ and ΔB are essentially independent and the optimum estimate of $\Delta\theta$ is f_2 and the optimum estimate of ΔB is $B - f_1$.

Furthermore the variance of ΔB is σ^2 and the variance of $\Delta\theta$ is $\sigma^2/Bf_1 \cong \sigma^2/B^2$. [illegible] estimate of velocity error is ω_c [illegible] $\left(1 - \frac{1}{B}\right) v_d(t)$ and the variance of the estimate is $\omega_c^2 \frac{\sigma^2}{Bf_1} x_d^2 + \frac{\sigma^2}{B^2} v_d^2 \cong \omega_c^2 \sigma^2$.

IV. Conclusion

A fairly general technique for the solution of the estimation problem has been presented. At the outset it was decided that suboptimum estimates would be acceptable because the resultant estimators would be easily realized in a physical system. The obvious question arises as to how the basis sets are chosen and how the nonlinear transformations are determined. The mathematician may be particularly critical of this aspect of the approach. The equally obvious defense is that the engineer is noted for his intuition in such problems and furthermore, he may be motivated to use his imagination and obtain solutions using this attack but give up before starting if required to solve a complicated integral equation. Many engineers would have obtained the results of Section III or something very similar. The flexibility and intrigue of the approach suggests that it could almost become a parlor game.

The one question that has not been answered is: To what extent are the solutions obtained suboptimum? That is, how much larger is the variance when compared to the true optimum and how much is the risk function increased. Other areas of further [illegible] [illegible] noise [illegible] the field is wide open [illegible] the application to other parameters when the observation is done in the presence of white Gaussian noise.

Related references, 8/6/73

65N90299   PR-2   NSG-712   65/00/00   4 PAGES   UNCLASSIFIED DOCUMENT
A RISK THEORY APPROACH TO NONLINEAR ESTIMATION PROBLEMS IN COMMUNICATION AND CONTROL PROGRESS REPORT, DEC. 16, 1964 - JUN. 16, 1965
A/PARK, J. H., JR.   <1965<   4 P   REFS
MINNESOTA UNIV., MINNEAPOLIS.   (DEPT. OF ELECTRICAL ENGINEERING.)

66X91169   NASA-CR-60265   NSG-712   64/00/00   6 PAGES   UNCLASSIFIED DOCUMENT   NASA
A RISK THEORY APPROACH TO NONLINEAR ESTIMATION PROBLEMS IN COMMUNICATION AND CONTROL   FIRST PROGRESS REPORT, 16 JUN. - 16 DEC. 1964
A/PARK, J. H., JR.
<1964<   6 P
MINNESOTA UNIV., MINNEAPOLIS.   (DEPT. OF ELECTRICAL ENGINEERING.)

70A29589*   ISSUE 13   PAGE 2367   CATEGORY 7   NSG-712   70/04/00   9 PAGES   UNCLASSIFIED DOCUMENT
AN FM DETECTOR FOR LOW S/N
(FM DETECTOR WITHOUT NORMALIZATION USEFUL FOR LOW S/N RECEPTION OF FSK AND PSK IN TELEMETRY SYSTEMS)
A/PARK, J. H., JR.   (AA/MINNESOTA, U., MINNEAPOLIS, MINN./.)
IEEE TRANSACTIONS ON COMMUNICATION TECHNOLOGY, VOL. COM-18, P. 110-118.
/*FREQUENCY MODULATION/*SIGNAL DETECTORS/*SIGNAL TO NOISE RATIOS/ AMPLITUDE MODULATION/ CORRELATORS/ FREQUENCY SHIFT KEYING/ PHASE SHIFT KEYING/ SINE WAVES/ SYSTEMS ENGINEERING/ TELEMETRY

68A19410   ISSUE 7   PAGE 1144   CATEGORY 7   68/01/00   2 PAGES   UNCLASSIFIED DOCUMENT
WHITE NOISE AND THE DELTA FUNCTION.
(WHITE NOISE AND DELTA FUNCTION WITH RESPECT TO SIGNAL SPACE, DISCUSSING AUTOCORRELATION FUNCTION)
A/PARK, J. H., JR.   (AA/MINNESOTA, U., DEPT. OF ELECTRICAL ENGINEERING, MINNEAPOLIS, MINN./.)
IEEE, PROCEEDINGS, VOL. 56, P. 114, 115.
/*AUTOCORRELATION/*DELTA FUNCTION/*SIGNAL TO NOISE RATIOS/*WHITE NOISE/ KERNEL FUNCTIONS

~~TYPE 4-2-6~~

66N29569*#   ISSUE 16   PAGE 3120   CATEGORY 19   NASA-CR-75892   NSG-712   65/00/00   22 PAGES   UNCLASSIFIED DOCUMENT
AN ENGINEERING APPROACH TO NONLINEAR ESTIMATION
(LINEAR ANALOG OPERATION FOLLOWED BY NONLINEAR DIGITAL SIGNAL PROCESSING FOR NONLINEAR PARAMETER APPROXIMATION)
A/PARK, J. H., JR.
MINNESOTA UNIV., MINNEAPOLIS.   AVAIL.NTIS
<1965<   22 P
/*APPROXIMATION METHOD/*COMPUTER METHOD/*DIGITAL SIMULATION/*NONLINEAR SYSTEM/ AMPLITUDE/ ANALOG/ COMPUTER/ DIGITAL/ FREQUENCY/ FUNCTION/ LINEAR/ METHOD/ MODULATION/ NONLINEAR/ PARAMETER/ SIGNAL/ SIMULATION/ SPACE/ SYSTEM